

AD-A125 130

STOCHASTIC CONVERGENCE PROPERTIES OF THE ADAPTIVE
GRADIENT LATTICE... PENNSYLVANIA STATE UNIV
UNIVERSITY PARK APPLIED RESEARCH LAB..

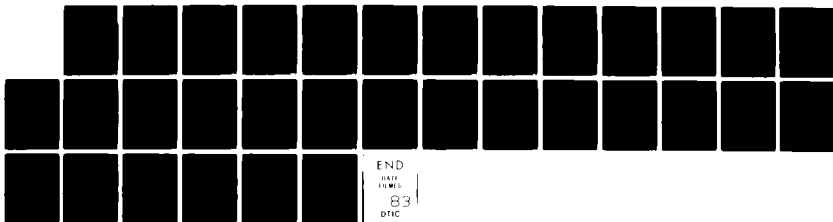
1/1

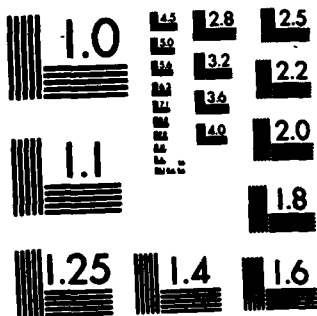
UNCLASSIFIED

G L SOHIE ET AL. 08 DEC 82

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A125130

6

STOCHASTIC CONVERGENCE PROPERTIES OF THE
ADAPTIVE GRADIENT LATTICE

G. R. L. Sohie and L. H. Sibul

Technical Memorandum
File No. TM 82-237
December 8, 1982
Contract No. N00024-79-C-6043

Copy No. 31

The Pennsylvania State University
Intercollege Research Programs and Facilities
APPLIED RESEARCH LABORATORY
Post Office Box 30
State College, PA 16801

Approved for Public Release
Distribution Unlimited

NAVY DEPARTMENT
NAVAL SEA SYSTEMS COMMAND

DTIC
ELECTE
MAR 1 1983
A

DTIC FILE COPY

03 03 001 100

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|--|-------------------------------|--|
| 1. REPORT NUMBER TM 82-237 | 2. GOVT ACCESSION NO. AD-A | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) STOCHASTIC CONVERGENCE PROPERTIES OF THE ADAPTIVE GRADIENT LATTICE | | 5. TYPE OF REPORT & PERIOD COVERED Interim |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) G. R. L. Sohie and L. H. Sibul | | 8. CONTRACT OR GRANT NUMBER(s) N00024-79-C-6043 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS Applied Research Laboratory P. O. Box 30 State College, PA 16801 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS Naval Sea Systems Command, Code 63R-14 Department of the Navy Washington, DC 20362 | | 12. REPORT DATE December 8, 1982 |
| | | 13. NUMBER OF PAGES 27 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | | 15. SECURITY CLASS. (of this report) Unclassified |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |
| 16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) | | |
| 18. SUPPLEMENTARY NOTES | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) stochastic, convergence, properties, adaptive, gradient, lattice | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) A stochastic fixed-point theorem is used as a basis for the study of stochastic convergence properties (in mean-squares sense) of the adaptive gradient lattice filter. Such properties include conditions on the stepsize in the adaptive algorithm and analytic expressions for the misadjustment and convergence rate. | | |

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Our results indicate that the limits on the stepsize are stricter than the ones obtained by considering convergence of the mean of the reflection coefficients and, therefore, only a slower convergence of the mean-square error can be obtained. It is shown that faster convergence is achieved for highly uncorrelated sequences (low S/N ratio) than for almost deterministic sequences (high S/N ratio). The misadjustment is shown to be exponentially dependent on the number of stages in the lattice and is higher for uncorrelated sequences than for almost deterministic sequences.

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)



| | |
|--------------------|-------------------------------------|
| Accession For | |
| NTIS GRA&I | <input checked="" type="checkbox"/> |
| DTIC TAB | <input type="checkbox"/> |
| Unannounced | <input type="checkbox"/> |
| Justification | |
| By | |
| Distribution/ | |
| Availability Codes | |
| Dist | Avail and/or Special |
| A | |

Subject: Stochastic Convergence Properties of the
Adaptive Gradient Lattice

References: See pages 22 and 23.

ABSTRACT

A stochastic fixed-point theorem is used as a basis for the study of stochastic convergence properties (in mean-squares sense) of the adaptive gradient lattice filter. Such properties include conditions on the stepsize in the adaptive algorithm and analytic expressions for the misadjustment and convergence rate.

Our results indicate that the limits on the stepsize are stricter than the ones obtained by considering convergence of the mean of the reflection coefficients and, therefore, only a slower convergence of the mean-square error can be obtained. It is shown that faster convergence is achieved for highly uncorrelated sequences (low S/N ratio) than for almost deterministic sequences (high S/N ratio). The misadjustment is shown to be exponentially dependent on the number of stages in the lattice and is higher for uncorrelated sequences than for almost deterministic sequences.

This paper has been accepted for presentation and publication in the Proceedings of 1983 IEEE International Conference on Acoustics, Speech and Signal Processing, Boston, MA, April 14, 15 and 16, 1983.

TABLE OF CONTENTS

| | <u>Page No.</u> |
|--|-----------------|
| ABSTRACT | 1 |
| TABLE OF CONTENTS | 2 |
| LIST OF FIGURES | 3 |
| I. INTRODUCTION | 4 |
| II. MATHEMATICAL BACKGROUND | 4 |
| 1. The Lattice Filter | 4 |
| 2. The Adaptive Lattice Filter | 9 |
| 3. The Stochastic Fixed-Point Theorem | 10 |
| III. CONVERGENCE PROPERTIES | 11 |
| 1. The Adaptive Time-Recursive Algorithm | 11 |
| 2. Conditions on the Stepsize (α) | 12 |
| 3. The Misadjustment | 15 |
| 4. The Convergence Rate | 19 |
| CONCLUSION | 21 |
| REFERENCES | 22 |
| APPENDIX A | 24 |
| APPENDIX B | 25 |

LIST OF FIGURES

| <u>Figure No.</u> | | <u>Page No.</u> |
|-------------------|---|-----------------|
| 1 | Lattice Filter as Cascade of Elementary Operators A_1 | 5 |
| 2 | i-th Element in the Lattice Filter | 6 |
| 3 | Model for the Total Output Misadjustment | 18 |

I. INTRODUCTION

Research results in various areas such as speech processing [1], array processing [2], and adaptive tracking [3] have indicated that the gradient-lattice algorithms have superior convergence behavior over classic time-recursive methods such as Widrow's LMS approach [4]. The majority of this work, however, has been of a primarily experimental nature. It is the purpose of this paper to provide a theoretical framework for the convergence study of such gradient-lattice algorithms. Since the convergence of adaptive algorithms is primarily a stochastic problem, a mean-square criterion will be considered as opposed to the convergence of the mean of the reflection coefficients, which is mostly used in the literature [1,3,5-10].

A stochastic version of a fixed-point theorem, which was developed by Oza [11-12], will provide the basis in establishing convergence conditions for the gradient-lattice algorithm. Such an approach has been taken in an earlier paper [13] to study convergence properties of the LMS algorithm and is based on the notion of a contraction mapping on a "stochastic" Hilbert space [13]. In the same framework, expressions for the convergence rate, conditions on the stepsize, and misadjustment are derived by considering a "distance" measure in the appropriate space.

II. MATHEMATICAL BACKGROUND

1. The Lattice Filter

The lattice filter, as depicted in Figure 1, will formally be considered as the cascade of elementary operators A_i (Fig. 2):

$$A = \prod_{i=1}^N A_i \quad (1)$$

December 8, 1982
GRLS:LHS:cac

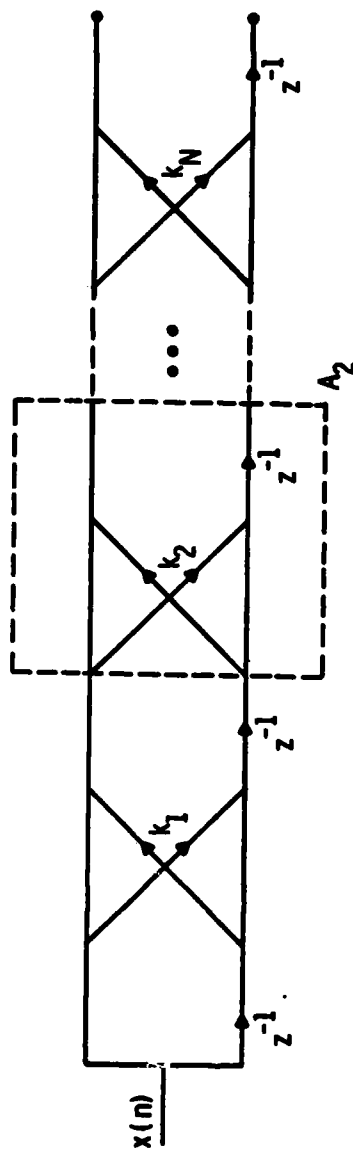


Figure 1. Lattice Filter as Cascade of Elementary Operators A_1 .

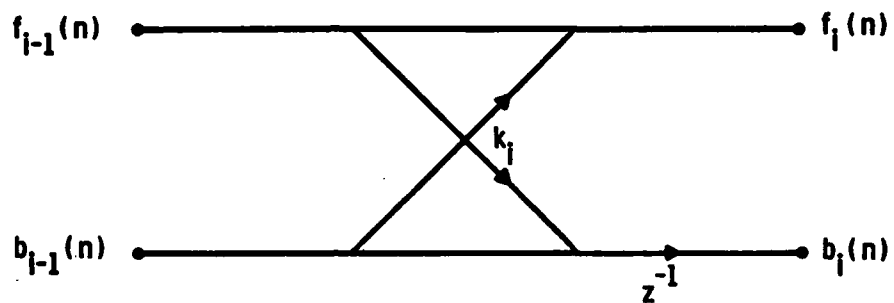


Figure 2. i -th Element in the Lattice Filter.

where N is the order of the filter. Since the input of each element in the lattice can be represented by the "augmented" vector:

$$\underline{e}_{i-1}(n) = \begin{bmatrix} f_{i-1}(n) \\ b_{i-1}(n) \end{bmatrix}, \quad (2)$$

A_i will be assumed to be an operator on $\ell_2(\Omega)$, the Hilbert space of zero-mean, wide-sense stationary stochastic (2x1) sequences with inner product:

$$\langle \underline{e}_i(\cdot), \underline{e}_j(\cdot) \rangle \triangleq E\{\underline{e}_j^T(n) \underline{e}_i(n)\} \quad (3)$$

It is also assumed that $\underline{e}_i(\cdot)$ has finite average power*. The correlation matrix at the input of each stage is given by:

$$E\{\underline{e}_i(n) \underline{e}_i^T(n)\} \triangleq \begin{bmatrix} E_i & C_i \\ C_i & E_i \end{bmatrix} \quad (4)$$

where, by symmetry, the average power of $f_i(n)$ and $b_i(n)$ (E_i) are equal.

A_i can be written in matrix form as:

$$\begin{aligned} \underline{e}_i(n) &= \begin{bmatrix} 1 & k_i z^{-1} \\ k_i & z^{-1} \end{bmatrix} \underline{e}_{i-1}(n) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \begin{bmatrix} 1 & k_i \\ k_i & 1 \end{bmatrix} \underline{e}_{i-1}(n) \triangleq U K_i \underline{e}_{i-1}(n) \end{aligned} \quad (5)$$

*The general theory developed in this paper is applicable to non-Gaussian stochastic processes. However, simple expressions for the appropriate norms can be obtained if Gaussian assumption is made.

where z^{-1} is the unit shift operator, and U and K_i are the matrices defined by:

$$U = \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \quad K_i = \begin{bmatrix} 1 & k_i \\ k_i & 1 \end{bmatrix} \quad (6)$$

where the magnitudes of the "reflection coefficients" k_i are always taken to be smaller than one. The norm of A_i [14] is given by:

$$\|A_i\| = \sup_{\|e_{i-1}(\cdot)\| = 1} \{\|U K_i e_{i-1}(\cdot)\|\} \quad (7)$$

where $\|\cdot\|$ denotes the norm induced by the inner product in (Eq. 3). Since it is straightforward to show that U is a unitary operator, (Eq. 7) reduces to the simple matrix norm of K_i , which is given by its spectral radius [14]:

$$\|A_i\| = |\lambda_{\max}| = 1 + |k_i| \quad (8a)$$

where λ_{\max} is the largest eigenvalue of A_i . A lower bound on the output power can be obtained by replacing \sup by \inf and using the same procedure. This gives:

$$(1 - |k_i|) \|e_{i-1}(\cdot)\| \leq \|e_i(\cdot)\| \leq (1 + |k_i|) \|e_{i-1}(\cdot)\| \quad (8b)$$

Note [10] that the lower bound is actually attained when $e_{i-1}(\cdot)$ consists of the actual innovations processes (optimal linear predictor). (Eq. 8b) was previously obtained by Makhoul using different methods [10].

2. The Adaptive Lattice Filter

Thus far, we have considered the lattice filter as a fixed, deterministic operator on the space $\ell_2(\Omega)$. If the lattice is used in an adaptive algorithm, however, two major differences arise. First of all, if a time-recursive method is used, the operators A_i will become time-varying and thus the assumption of wide-sense stationarity of the input is no longer valid. In the sequel, however, we will assume that the stepsize in the adaptive algorithm is sufficiently small such that the inputs to successive stages are at least locally stationary. Secondly, the parameters determining the operators A_i will be based on measured data, thus resulting in a stochastic operator. Consequently, the assumption of Gaussian statistics of the input to each stage will be violated. Again, under the same assumption of a small stepsize, it can be assumed that the statistics of the sequences $e_i(\cdot)$ will be sufficiently Gaussian.

In order to calculate the mean-square norm of the stochastic lattice element, we can again use (Eq. 5) where now the reflection coefficients k_i in the matrix K_i are random variables:

$$\begin{aligned} \|A_i\|^2 &= \sup_{\|e_{i-1}(\cdot)\| = 1} \{ \|K_i e_{i-1}(\cdot)\|^2 \} \\ &= \sup_{\|e_{i-1}(\cdot)\| = 1} E\{ e_{i-1}^T(n) K_i^2 e_{i-1}(n) \} \end{aligned} \quad (9)$$

Under the widely used assumption that the operator K_i and the sequence $e_{i-1}(\cdot)$ are uncorrelated, (Eq. 9) becomes:

$$\begin{aligned} \|A_i\|^2 &= \sup_{\|e_{i-1}(\cdot)\| = 1} E\{e_{i-1}^T(n) E\{K_i^2\} e_{i-1}(n)\} \\ &\leq \|E\{K_i^2\}\| \end{aligned} \quad (10)$$

(by the Schwarz inequality and the definition of matrix norm).

Because the matrix norm is attained by an eigenvector and since in this case the equality sign in the Schwarz inequality holds, we have:

$$\|A_i\|^2 = \|E\{K_i^2\}\| \quad (11)$$

Noting that:

$$\begin{aligned} E\{K_i^2\} &= E \left\{ \begin{bmatrix} 1+k_i^2 & 2k_i \\ 2k_i & 1+k_i^2 \end{bmatrix} \right\} \\ &= \begin{bmatrix} 1+\sigma_i^2+\bar{k}_i^2 & 2\bar{k}_i \\ 2\bar{k}_i & 1+\sigma_i^2+\bar{k}_i^2 \end{bmatrix} \end{aligned} \quad (12)$$

where σ_i^2 and \bar{k}_i are the variance and mean, respectively, of k_i , we can write bounds for $\|e_i(\cdot)\|$ similarly to (Eq. 8b):

$$[\sigma_i^2 + (1-|\bar{k}_i|)^2] \|e_{i-1}(\cdot)\|^2 \leq \|e_i(\cdot)\|^2 \leq [\sigma_i^2 + (1+|\bar{k}_i|)^2] \|e_{i-1}(\cdot)\|^2 \quad (13)$$

3. The Stochastic Fixed-Point Theorem

The following theorem, which forms the basis for the discussions in the sequel, is a stochastic fixed-point theorem and was used by Oza [11,12] in a system identification problem. For the proof, we refer to [11].

Theorem

Let $\{T_n\}_{n=0}^{\infty}$ be a sequence of (random) operators on a Hilbert space H and let $T_n \rightarrow T$ where T is a contraction mapping, i.e.

$$\lim_{n \rightarrow \infty} \|T_n y - T y\| = 0 \quad \forall y \in H \quad (14)$$

and

$$\|T y_1 - T y_2\| < \|y_1 - y_2\| \quad \forall y_1, y_2 \in H \quad (15)$$

Then the sequence generated by:

$$y_{n+1} = T_n y_n$$

y_0 fixed, but arbitrary in H , converges strongly to the fixed-point of T .

III. CONVERGENCE PROPERTIES

1. The Adaptive Time-Recursive Algorithm

It is well-known [15] that the lattice filter is a natural implementation of the Levinson algorithm to solve the linear prediction problem. In the "optimal" case, the reflection coefficients k_i are determined in the Levinson recursion by:

$$E_{i-1} k_i = -C_{i-1} \quad (16)$$

and the sequences $f_i(\cdot)$ and $b_i(\cdot)$ are the innovations processes of the i th-order forward and backward prediction [16]. Practically, however, neither the average power (E_i) nor the crosscorrelation (C_i) of the innovations processes are completely known, and an adaptive method then consists of calculating the reflection coefficients based on estimates

of C_i and E_i which are obtained from actual data. A very common and simple form of adaptation is achieved when the average values in (Eq. 4) are replaced by their instantaneous values and the solutions to (Eq. 16) are obtained time-recursively. This leads to an adaptive algorithm given by (fr Ref. [3]):

$$k_{i,n+1} = k_{i,n} - \alpha [\hat{E}_{i-1}(n) k_{i,n} + \hat{C}_{i-1}(n)] \quad (17)$$

where

$$\hat{E}_{i-1}(n) \triangleq \frac{1}{2} [f_{i-1}^2(n) + b_{i-1}^2(n)] \quad (18)$$

and

$$\hat{C}_{i-1}(n) \triangleq f_{i-1}(n) b_{i-1}(n) \quad (19)$$

Note that if the estimates are taken to be the exact values (i.e., $\hat{E}_{i-1}(n) \equiv E_{i-1}$, $\hat{C}_{i-1}(n) \equiv C_{i-1}$), the sequence in (Eq. 17) converges if $\alpha < \frac{2}{E_{i-1}}$, the same condition as is obtained for convergence of the mean of the reflection coefficients [3].

2. Conditions on the Step size (α)

Strictly speaking, an adaptive lattice element $A_{i,n}$ converges to a lattice element A_i (the "optimal" operator) if:

$$\lim_{n \rightarrow \infty} \|A_{i,n} e_{i-1}(\cdot) - A_i e_{i-1}(\cdot)\| = 0 \quad (20)$$

Practically, however, most adaptive algorithms can only result in outputs which are within a certain "distance" of the optimal value, i.e.,

$$\lim_{n \rightarrow \infty} \|A_{i,n} e_{i-1}(\cdot) - A_i e_{i-1}(\cdot)\| = M_i < \infty \quad (21)$$

where M_i is the (unnormalized) misadjustment. Using (Eq. 5), the norm in (Eq. 21) can be written as:

$$\| (A_{i,n} - A_i) \varepsilon_{i-1}(\cdot) \|^2 = \left\| \begin{bmatrix} 0 & k_{i,n} - k_i \\ k_{i,n} - k_i & 0 \end{bmatrix} \varepsilon_{i-1}(\cdot) \right\|^2 \quad (22)$$

$$= E\{|k_{i,n} - k_i|^2\} E_{i-1} \quad (23)$$

Thus, convergence of the adaptive element is determined by convergence of the reflection coefficients.

In order to obtain limits for the stepsize α , (Eq. 17) can be written in terms of a stochastic fixed-point theorem as follows: write (Eq. 17) as:

$$\begin{aligned} k_{i,n+1} - k_i &= (k_{i,n} - k_i) [1 - \alpha \hat{E}_{i-1}(n)] \\ &\quad - \alpha [\hat{C}_{i-1}(n) + k_i \hat{E}_{i-1}(n)] \end{aligned} \quad (24)$$

and define the stochastic operator $\hat{T}_{i,n}$ on the space of Gaussian random variables by:

$$\hat{T}_{i,n} k \triangleq [1 - \alpha \hat{E}_{i-1}(n)] k - \alpha [\hat{C}_{i-1}(n) + k_i \hat{E}_{i-1}(n)] \quad (25)$$

for all k . Then, (Eq. 24) can be expressed as a fixed-point problem:

$$(k_{i,n+1} - k_i) = \hat{T}_{i,n} (k_{i,n} - k_i) \quad (26)$$

and using the contraction mapping principle of Section I.3, the recursion in (Eq. 26) will converge if:

$$\| [1 - \alpha \hat{E}_{i-1}(n)] (y_1 - y_2) \| < \| y_1 - y_2 \| \quad (27)$$

for y_1 and y_2 fixed but arbitrary.

Since the norm in (Eq. 27) is given by:

$$\begin{aligned} & \| [1 - \alpha \hat{E}_{i-1}(n)](y_1 - y_2) \|^2 \\ &= E\{|1 - \alpha \hat{E}_{i-1}(n)|^2\} \|y_1 - y_2\|^2 \end{aligned} \quad (28)$$

this condition becomes (see Appendix A):

$$1 + \alpha^2 [2 E_{i-1}^2 + C_{i-1}^2] - 2 \alpha E_{i-1} < 1 \quad (29)$$

or

$$0 < \alpha < \frac{2E_{i-1}}{2E_{i-1}^2 + C_{i-1}^2} \quad (30)$$

Various interpretations can now be made. First of all, it is easy to see from (Eq. 30) that the limits obtained in this fashion are stricter than the ones for convergence of the mean of $k_{i,n}$. Furthermore, while previous results showed a dependence of the upper limit on E_{i-1} only, and thus the upper limit for α could be made independent of the position in the lattice by normalizing it with respect to E_{i-1} [3], Equation (30) shows that α is also dependent on the crosscovariance of $f_{i-1}(\cdot)$ and $b_{i-1}(\cdot)$.

Thus, in order to make the limits for the stepsize independent on the position in the ladder, normalizing with respect to the right-hand side of (Eq. 30) requires the extra computation of C_{i-1} . As C_{i-1} is bounded by the Schwarz inequality as:

$$0 \leq |C_{i-1}| \leq E_{i-1} \quad (31)$$

two extreme cases can be considered. If $|C_{i-1}|$ is very small (input

sequence extremely random) the upper bound for α in (Eq. 30) becomes approximately equal to $\frac{1}{E_{i-1}}$, while if the input sequence is almost deterministic (singular process), the limit for α approaches $\frac{2}{3} \frac{1}{E_{i-1}}$. This indicates that faster convergence can be obtained for lower signal-to-noise ratios, a result which seems to be supported by some experimental work.

3. The Misadjustment

The misadjustment due to the i th stage in the lattice filter is defined by (Eq. 21). Since (Eq. 23) shows that

$$M_i^2 = \lim_{n \rightarrow \infty} E\{|k_{i,n} - k_i|^2\} E_{i-1} \quad (32)$$

it is clear that the normalized misadjustment is nothing but the norm in the fixed-point problem (Eq. 26):

$$m_i^2 \triangleq \frac{M_i^2}{E_{i-1}} = \lim_{n \rightarrow \infty} \|\hat{T}_{i,n}(k_{i,n} - k_i)\| \quad (33)$$

The remark can be made that the fixed-point of T_i , the "ideal" contraction mapping given by

$$T_i k \triangleq [1 - \alpha E_{i-1}]k - \alpha[C_{i-1} + k_i E_{i-1}] \quad (34)$$

has as (only) fixed-point the zero element because

$$\begin{aligned} T_i 0 &= 0 - \alpha[C_{i-1} + k_i E_{i-1}] \\ &= 0 \end{aligned} \quad (35)$$

by the definition of k_i (Eq. 16). Thus, zero misadjustment can be obtained if the estimates in $\hat{E}_{i-1}(\cdot)$ and $\hat{C}_{i-1}(\cdot)$ converge to the actual values.

This will be the case if a different estimate is used as:

$$\hat{E}_{i-1}(n) = \frac{1}{2n} \sum_{j=1}^n [f_{i-1}^2(j) + b_{i-1}^2(j)] \quad (36)$$

and

$$\hat{C}_{i-1}(n) = \frac{1}{n} \sum_{j=1}^n [f_{i-1}(j) b_{i-1}(j)] \quad (37)$$

and if the process $\underline{e}(\cdot)$ is correlation ergodic [18]. Note that in this case, $\frac{\alpha}{n}$ satisfies the established conditions for stochastic approximation [19].

In order to compute the misadjustment of the algorithm (Eq. 17), the norm of (Eq. 26) is needed.

$$\|k_{i,n+1} - k_i\|^2 = \|\hat{T}_{i,n}(k_{i,n} - k_i)\|^2 \quad (38)$$

This norm is computed in Appendix B and the (normalized) misadjustment is found to be:

$$m_i^2 = \alpha \left[\frac{E_{i-1}^2 - C_{i-1}^2}{E_{i-1}} \right]^2 \cdot \frac{1}{2E_{i-1} - \alpha[2E_{i-1}^2 + C_{i-1}^2]} \quad (39)$$

This expression indicates that, while the limits on α can be made independent of the stage i by appropriate normalization of the stepsize, such normalization will not result in a stage-independent misadjustment. Again, two extreme cases occur. If the input sequence is extremely random ($|C_{i-1}| \ll E_{i-1}$), (Eq. 39) will equal approximately $\frac{\alpha E_{i-1}}{2(1 - \alpha E_{i-1})}$, while if the input sequence is almost deterministic, the misadjustment approaches zero. Therefore, M_i is bounded by:

$$0 < M_i^2 < \frac{\alpha E_{i-1}^2}{2(1 - \alpha E_{i-1})} \quad (40)$$

In this result, E_{i-1} is the average power in steady state conditions at the input of the i th stage. For the usual case where α is normalized with respect to E_{i-1} , the upper bound in (Eq. 40) becomes

$$0 < M_i^2 < \frac{\alpha}{2(1 - \alpha)} E_{i-1} \quad (41)$$

In order to compute the total misadjustment at the output of the cascade, (Eq. 13) can be used. Essentially, (Eq. 13) corresponds to the model shown in Figure 3. The total output power in Figure 3 is bounded by (Eq. 13) where σ_i^2 and \bar{k}_i are (since the lattice converged) equal to m_i^2 and k_i , respectively:

$$\sum_{i=1}^N [m_i^2 + (1 - |k_i|)^2] E_o \leq E_N \leq \sum_{i=1}^N [m_i^2 + (1 + |k_i|)^2] E_o \quad (42)$$

Since the optimal output power is given in (Eq. 8b), the output misadjustment normalized with respect to the total output power is given by:

$$\left\{ \sum_{i=1}^N \left[\frac{m_i^2 + (1 - |k_i|)^2}{(1 - |k_i|)^2} \right] \right\} - 1 \leq m \leq \left\{ \sum_{i=1}^N \left[\frac{m_i^2 + (1 + |k_i|)^2}{(1 - |k_i|)^2} \right] \right\} - 1 \quad (43)$$

If the input sequence is very random, $|k_i| \ll 1$ and both bounds approach the same value. In this case, we will have approximately

$$m \approx \sum_{i=1}^N (1 + m_i^2) - 1 \quad (44)$$

If α is normalized with respect to E_{i-1} , which is the case in most applications, we have from (Eq. 41):

December 8, 1982
GRLS:LHS:cac

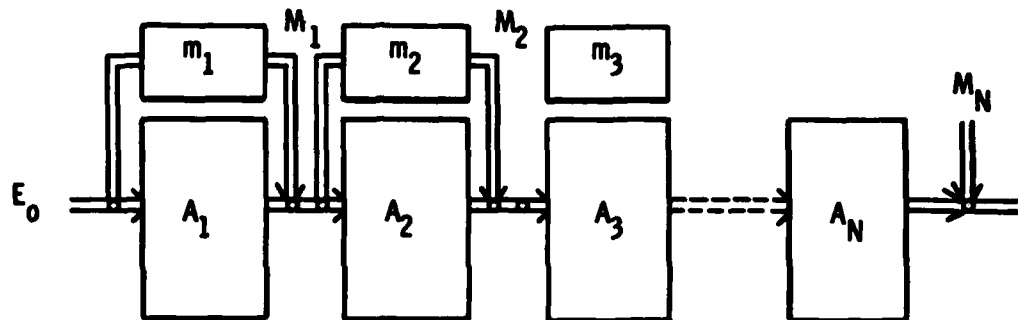


Figure 3. Model for the Total Output Misadjustment.

$$m_i^2 = \frac{\alpha}{2(1-\alpha)} \quad (45)$$

and finally

$$m = \left[\frac{2-\alpha}{2(1-\alpha)} \right]^N - 1 \quad (46)$$

Ungerboeck [21] and Widrow [22] showed that for the LMS algorithm, the misadjustment is linearly dependent on the order N. The exponential dependence on N for the lattice suggests that, in general, a higher misadjustment can be expected for the lattice filter.

4. The Convergence Rate

In order to compute the convergence rate of the *i*th adaptive element, the homogeneous part of the state equation in (Appendix B - Eq. B8) has to be considered. This is denoted by:

$$\underline{x}(n+1) = \begin{bmatrix} A & -2\alpha C \\ 0 & 1 - \alpha E_{i-1} \end{bmatrix} \underline{x}(n) \quad (47)$$

$$= M \underline{x}(n)$$

The convergence rate is determined by the smallest eigenvalue of the matrix *M* [20]. Since it is straightforward that this smallest eigenvalue is equal to *A*, the "slowest" mode, which determines convergence of the squared-error, is given by:

$$m_{i,n} = A^n \quad (48)$$

Substituting for *A* using (Appendix A - Eq. A3) and assuming small α , an approximate expression for $m_{i,n}$ is obtained:

$$m_{i,n} = 1 - n\{2\alpha E_{i-1} - \alpha^2[2E_{i-1}^2 + C_{i-1}^2]\} \quad (49)$$

Comparing (Eq. 49) to the expression:

$$e^{-t/\tau} = 1 - \frac{t}{\tau} + \dots$$

the time-constant of the adaptation is given by:

$$\tau_i = \frac{1}{2\alpha E_{i-1} - \alpha^2[2E_{i-1}^2 + C_{i-1}^2]} \quad (50)$$

In the extreme case of almost deterministic inputs (high signal-to-noise ratio), this becomes:

$$\tau_i \approx \frac{1}{2\alpha E_{i-1}[1 - \frac{3}{2}\alpha E_{i-1}]} \quad (51)$$

and in case of highly uncorrelated input sequences (low S/N ratio), τ_i becomes:

$$\tau_i \approx \frac{1}{2\alpha E_{i-1}[1 - \alpha E_{i-1}]} \quad (52)$$

Thus, a slower adaptation is obtained for deterministic sequences than for uncorrelated sequences.

CONCLUSION

Our results, which are based on a mean-square convergence criterion, suggest that the usual approach, which consists of the study of convergence properties of the mean of the reflection coefficients, may not be sufficient for a number of practical applications. The mean-square approach shows that convergence is only obtained under stricter conditions than previous studies suggested. The convergence rate of the adaptive gradient lattice is better for highly uncorrelated input signals (low signal-to-noise ratio) than for almost deterministic sequences (high signal-to-noise ratio), but a higher misadjustment can be expected if the signal-to-noise ratio is low. The total output misadjustment varies exponentially with the filter order. Since the output misadjustment for the basic LMS algorithm varies linearly with the number of taps, the excess mean-square error for the lattice can be expected to be larger, especially for higher order filters.

REFERENCES

- [1] J. Makhoul, "Stable and Efficient Lattice Methods for Linear Prediction," IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-25(5), pp. 423-428, October 1977.
- [2] W. F. Gabriel, "Spectral Analysis and Adaptive Array Super-resolution Techniques," IEEE Proceedings, Vol. 68(6), pp. 654-666, June 1980.
- [3] W. S. Hodgkiss, J. A. Presley, Jr., "Adaptive Tracking of Multiple Sinusoids whose Power Levels are Widely Separated," IEEE Trans. Acoustics, Speech and Signal Processing, Vol. 29(3), pp. 710-721, June 1981.
- [4] B. Widrow, "Adaptive Filters" in Aspects of Network and System Theory, R. Kalman and N. DeClaris, eds., pp. 563-587, Holt, Rinehart and Winston, 1971, pp. 563-587.
- [5] J. Makhoul, R. Visuanathan, "Adaptive Lattice Methods for Linear Prediction," Proc. International Conference on Acoustics, Speech and Signal Processing, pp. 83-86, April 1978.
- [6] R. Visuanathan, J. Makhoul, "Sequential Lattice Methods for Stable Linear Prediction," Eascon '76, pp. 155A - 155H.
- [7] L. J. Griffiths, "Adaptive Structures for Multiple - Input Noise Cancelling Applications," Proc. International Conference on Acoustics, Speech and Signal Processing, pp. 925-928, April 2-4, 1979.
- [8] L. J. Griffiths, "An Adaptive Lattice Structure for Noise-Cancelling Applications," Proc. International Conference on Acoustics, Speech and Signal Processing, pp. 87-90, April 10-12, 1978.
- [9] L. J. Griffiths, "A Continuously - Adaptive Filter as a Lattice Structure," Proc. International Conference on Acoustics, Speech and Signal Processing, pp. 683-686, May 13-19, 1977.
- [10] J. Makhoul, "A Class of All-Zero Lattice Digital Filters, Properties and Applications," IEEE Trans. Acoustics, Speech and Signal Processing, ASSP-26(6), pp. 304-314, August 1978.

- [11] K. G. Oza, "Identification Problem and Random Contraction Mappings," Ph.D. Dissertation, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, 1967.
- [12] K. G. Oza, E. I. Jury, "System Identification and the Principle of Random Contraction Mapping," SIAM Journal on Control, Vol. 6(2), pp. 249-257, 1968.
- [13] G. R. L. Sohie, L. H. Sibul, "Application of Hilbert Space Theory to Optimal and Adaptive Array Processing," 16th Annual Conference on Information Sciences and Systems, Princeton University, March 17-19, 1982.
- [14] F. Riesz, B. Sz-Nagy, Functional Analysis, Translated by F. Boron, F. Ungar Publishing Co., New York, 1955.
- [15] J. D. Markel, A. H. Gray, Linear Prediction of Speech, New York, Springer-Verlag, 1979.
- [16] T. Kailath, "A View of Three Decades of Linear Filtering Theory," IEEE Trans. on Information Theory, Vol. IT-20(2), pp. 146-181, March 1974.
- [17] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill Book Company, 1965.
- [18] A. Papoulis, Signal Analysis, McGraw-Hill, New York, 1977.
- [19] J. Sakrison, "Stochastic Approximation: A Recursive Method for Solving Regression Problems," in Advances in Communication Systems, A. V. Balakrishnam, editor, Academic Press, 1966, New York & London.
- [20] H. Kavakernaak, R. Sivan, Linear Optimal Control Systems, Wiley-Interscience, New York, 1972.
- [21] G. Ungerboeck, "Theory on the Speed of Convergence in Adaptive Equalizers for Digital Communication," IBM J. Res. Development, November 1972.
- [22] B. Widrow et al., "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter," IEEE Proceedings, Vol. 64(8), pp. 1151-1162, August 1976.

APPENDIX A

Note that:

$$\begin{aligned}
 & E\{|1 - \alpha \hat{E}_{i-1}(n)|^2\} \\
 &= E\{1 + \alpha^2 |\hat{E}_{i-1}(n)|^2 - 2\alpha \hat{E}_{i-1}(n)\} \\
 &= 1 + \alpha^2 E\{|\hat{E}_{i-1}(n)|^2\} - 2\alpha E_{i-1}
 \end{aligned} \tag{A1}$$

Now, using the well-known expression for 4-th order Gaussian moments [17], this becomes from (Eq. 18):

$$\begin{aligned}
 E\{|\hat{E}_{i-1}(n)|^2\} &= \frac{1}{4} E\{|f_{i-1}^2(n) + b_{i-1}^2(n)|^2\} \\
 &= \frac{1}{4} \left[E\{f_{i-1}^4(n) + b_{i-1}^4(n)\} \right. \\
 &\quad \left. + 2 E\{f_{i-1}^2(n) b_{i-1}^2(n)\} \right] \\
 &= \frac{1}{4} \left[3 E_{i-1}^2 + 3 E_{i-1}^2 + 2 E_{i-1}^2 + 4 C_{i-1}^2 \right] \\
 &= 2 E_{i-1}^2 + C_{i-1}^2
 \end{aligned} \tag{A2}$$

Substituting this in (Appendix A - Eq. A1), we finally obtain:

$$E\{|1 - \alpha \hat{E}_{i-1}(n)|^2\} = \alpha^2 [2 E_{i-1}^2 + C_{i-1}^2] + 1 - 2\alpha E_{i-1} \tag{A3}$$

APPENDIX B

Using (Eq. 25), the following recursion can be written:

$$\begin{aligned}
 \|k_{i,n+1} - k_i\|^2 &= E\{|[1 - \alpha \hat{E}_{i-1}(n)](k_{i,n} - k_i) - \alpha[\hat{C}_{i-1}(n) + k_i \hat{E}_{i-1}(n)]|^2\} \\
 &= E\{|[1 - \alpha \hat{E}_{i-1}(n)]^2 (k_{i,n} - k_i)^2\} \\
 &\quad + \alpha^2 E\{[\hat{C}_{i-1}(n) + k_i \hat{E}_{i-1}(n)]^2\} \\
 &\quad - 2\alpha E\{[1 - \alpha \hat{E}_{i-1}(n)] [\hat{C}_{i-1}(n) + k_i \hat{E}_{i-1}(n)] (k_{i,n} - k_i)\} \quad (B1)
 \end{aligned}$$

If we assume that the repetition rate of the algorithm is small enough such that $(k_{i,n} - k_i)$ is uncorrelated with $\hat{C}_{i-1}(n)$ and $\hat{E}_{i-1}(n)$, (Appendix B - Eq. B1) will become:

$$\begin{aligned}
 \|k_{i,n+1} - k_i\|^2 &= E\{|1 - \alpha \hat{E}_{i-1}(n)|^2\} \|k_{i,n} - k_i\|^2 \\
 &\quad + \alpha^2 E\{[\hat{C}_{i-1}(n) + k_i \hat{E}_{i-1}(n)]^2\} \\
 &\quad - 2\alpha E\{[1 - \alpha \hat{E}_{i-1}(n)] [\hat{C}_{i-1}(n) + k_i \hat{E}_{i-1}(n)]\} \\
 &\quad E\{k_{i,n} - k_i\} \quad (B2)
 \end{aligned}$$

The coefficients in (Appendix B - Eq. B2) can again be computed using the expression for 4-th order moments of Gaussian processes:

$$\begin{aligned}
 &E\{[\hat{C}_{i-1}(n) + k_i \hat{E}_{i-1}(n)]^2\} \\
 &= E\{|\hat{C}_{i-1}(n)|^2\} + k_i^2 E\{|\hat{E}_{i-1}(n)|^2\} + 2k_i E\{\hat{C}_{i-1}(n) \hat{E}_{i-1}(n)\} \quad (B3)
 \end{aligned}$$

In (Appendix B - Eq. B3) we have:

$$\begin{aligned} E\{|\hat{C}_{i-1}(n)|^2\} &= E\{f_{i-1}^2(n) b_{i-1}^2(n)\} \\ &= E_{i-1}^2 + 2 C_{i-1}^2 \end{aligned} \quad (B4)$$

$$\begin{aligned} E\{\hat{C}_{i-1}(n) \hat{E}_{i-1}(n)\} &= \frac{1}{2} E\{f_{i-1}(n) b_{i-1}(n) f_{i-1}^2(n)\} \\ &\quad + \frac{1}{2} E\{f_{i-1}(n) b_{i-1}(n) b_{i-1}(n)\} \\ &= 3 C_{i-1} E_{i-1} \end{aligned} \quad (B5)$$

Using (Eq. B4) and (Eq. B5) and (Eq. A2), (Eq. B3) becomes:

$$\begin{aligned} E\{[\hat{C}_{i-1}(n) + k_i \hat{E}_{i-1}(n)]^2\} \\ &= E_{i-1}^2 + 2 C_{i-1}^2 + 2 k_i^2 E_{i-1}^2 + k_i^2 C_{i-1}^2 \\ &\quad + 6 k_i C_{i-1} E_{i-1} \end{aligned} \quad (B6)$$

and with the value for the optimal reflection coefficient from (Eq. 16),

(Appendix B - Eq. B6) reduces to:

$$\begin{aligned} E\{[\hat{C}_{i-1}(n) + k_i \hat{E}_{i-1}(n)]^2\} &= E_{i-1}^2 + 2 C_{i-1}^2 + 2 C_{i-1}^2 + \frac{C_{i-1}^4}{E_{i-1}^2} - 6 C_{i-1}^2 \\ &= E_{i-1}^2 - 2 C_{i-1}^2 + \frac{C_{i-1}^4}{E_{i-1}^2} \\ &= \left[\frac{E_{i-1}^2 - C_{i-1}^2}{E_{i-1}} \right]^2 \end{aligned} \quad (B7)$$

The recursion in (Eq. B2) can therefore be written as:

$$\begin{aligned} \|k_{i,n+1} - k_i\|^2 = & A \|k_{i,n} - k_i\|^2 + \alpha^2 B \\ & - 2 \alpha C E\{k_{i,n} - k_i\} \end{aligned} \quad (B8)$$

where A is given by (Eq. A2), and B is given by (Eq. B7). Equation (B8) can now be considered as a state equation with state variable given by:

$$\underline{x}(n) \triangleq \begin{bmatrix} \|k_{i,n} - k_i\|^2 \\ E\{k_{i,n} - k_i\} \end{bmatrix} \quad (B9)$$

Transfer analysis of this state equation [20] and application of the final-value theorem of the Z-transform lead to an expression for the (normalized) misadjustment due to the *i*th stage:

$$m_i^2 = \frac{\alpha^2 B}{1-A} \quad (B10)$$

DISTRIBUTION LIST FOR ARL TM 82-237, By G. R. L. Sohie and L. H. Sibul,
dated December 8, 1982

Commander
Naval Sea Systems Command
Department of the Navy
Washington, DC 20362

Attention: Dr. E. G. Liszka, PMS 406B
Copy No. 1

Attention: Mr. D. Porter, SEA 63R1
Copy No. 2

Attention: Mr. D. C. Houser, SEA 63R-14
Copy No. 3

Attention: Mr. C. D. Smith, SEA 63R
Copy No. 4

Attention: Mr. J. D. Antinucci, PMS 402
Copy No. 5

Attention: CAPT R. M. Wellborn, PMS 406
Copy No. 6

Attention: Mr. T. E. Douglas, PMS 406E1
Copy No. 7

Attention: Code SEA 9961 (Library)
Copies 8 and 9

Commanding Officer
Naval Underwater Systems Center
Newport, RI 02840

Attention: Dr. J. R. Short
Copy No. 10

Attention: Dr. J. M. D'Albora
Copy No. 11

Commander
Naval Air Systems Command
Department of the Navy
Washington, DC 20362

Attention: Mr. J. G. Smith
Copy No. 12

Director
Applied Research Laboratory
University of Texas
Austin, TX 78712

Attention: Dr. J. F. Willman
Copy No. 13

Marine Physical Laboratory
Scripps Institute of Oceanography
San Diego, CA 92152

Attention: W. S. Hodgkiss
Copy No. 14

Attention: K. M. Watson
Copy No. 15

Director
Applied Physics Laboratory
University of Washington
Seattle, WA 98105

Attention: Mr. C. Sienkiewicz
Copy No. 16

DARPA
525 Architect Building
1400 Wilson Boulevard
Arlington, VA 22209

Attention: Dr. Theodore Kooij
Copy No. 17

Commanding Officer
Naval Coastal Systems Center
Panama City, FL 32401

Attention: Dr. David Skinner
Copy No. 18

Bolt Beranek and Newman, Inc.
1300 North 17th Street
Arlington, VA 22209

Attention: Dr. Henry Cox
Copy No. 19

Commander
Office of Naval Research
Department of the Navy
800 N. Quincy Street
Arlington, VA 22217

Defense Technical Information Center
Cameron Station
Alexandria, VA 22314

Copies 26 through 31

Attention: Statistics and
Probability Program
Code 411SP
Copy No. 20

Department of Electrical Engineering
and Computer Science
Princeton University
Princeton, NJ 08450

Attention: Prof. Stuart C. Schwartz
Copy No. 21

Department of Electrical Engineering
Kelly Hall
University of Rhode Island
Kingston, RI 02881

Attention: Prof. Donald W. Tufts
Copy No. 22

ORNICON Corporation
3366 N. Torrey Pines Ct., Suite 320
La Jolla, CA 92037

Attention: Dr. Richard A. Altes
Copy No. 23

Systems Control Technology, Inc.
1801 Page Mill Road
P. O. Box 10180
Palo Alto, CA 94303

Attention: Dr. Benjamin Friedlander
Copy No. 24

Department of Electrical Engineering
College of Engineering and Applied Science
The University of Rochester
Rochester, NY 14627

Attention: Prof. E. L. Titlebaum
Copy No. 25